Creativity Assessment in Neuroscience Research

Mathias Benedek[1,*], Alexander P. Christensen[2], Andreas Fink[1], & Roger E. Beaty[3]

1 Institute of Psychology, University of Graz, Austria; BioTechMed Graz

2 Department of Psychology, University of North Carolina at Greensboro, USA

3 Department of Psychology, Penn State University, USA

Accepted in *Psychology of Aesthetics, Creativity, and the Arts*

Special Issue: “*Creativity Assessment - Pitfalls, Solutions, and Standards*”

*Corresponding author:

Dr. Mathias Benedek, Institute of Psychology, University of Graz. Universitätsplatz 2, 8010 Graz, Austria. Phone: +43 316 380 8475. E-mail: mathias.benedek@uni-graz.at.

Cite as: Benedek, M., Christensen, A. P., Fink, A., & Beaty, R. E. (in press). Creativity assessment in neuroscience research. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. doi: 10.1037/aca0000215

Abstract

The investigation of the neural correlates of creative cognition requires researchers to adapt creativity tasks to meet the constraints imposed by cognitive neuroscience research – assessing well-defined cognitive processes, repeated over many tasks. We present a brief review of essential study design parameters in neuroscience research on creativity including number of task repetitions (i.e., trials), time on task, what kind of responses are collected (e.g., whether participants speak, write, draw, press buttons etc.), and when these responses are collected (e.g., after or during task). We further examine how design parameters depend on neuroscience methods (e.g., fMRI, EEG) and task type (e.g., divergent thinking, creative problem solving). The review discloses a substantial heterogeneity of methodological approaches across studies, but also identifies some established common practices. Typical adaptations include the employment of shortened tasks, which allows the realization of more tasks per session, and a more focused investigation of time-critical cognitive processes. Study designs also commonly separate periods of creative thought from response production in order to restrict the effect of response-related motor artifacts and to assess brain activity unique to the generation of creative ideas/solutions. We discuss the pros and cons of the various approaches with respect to the goal to increase reliability of neurophysiological measurements while maintaining valid assessments, and derive some recommendations for future research.

Keywords: Assessment; Creativity; Creative Cognition; Brain; EEG; fMRI

Neuroscience provides powerful methods to investigate cognitive processes in creative thought. In the last two decades, many scholars have joined the venture to unveil how creativity emerges in the brain. A special issue in the journal *Methods* supported these efforts by introducing novel approaches to the study neuroscience of creativity (Dietrich, 2007). Early reviews of research findings, however, observed little consistency in how creativity is manifested in the brain, which was partly attributed to the large heterogeneity of tasks and methods employed in the field (Arden et al., 2010; Dietrich et al., 2010; Sawyer, 2011). Indeed, studies have investigated very diverse creative activities ranging from drawing to musical improvisation to idea generation (Abraham, 2018). Moreover, even studies focusing on the same experimental task (e.g. alternate uses test) differed considerably in the specific way the task was implemented in the neurophysiological assessment (e.g., task duration, mode of responding, control tasks, etc.). Later reviews that put a focus on more specific creative activities and brain imaging methods provided more consistent results (e.g., Boccia et al., 2015; Fink & Benedek, 2014; Gonen-Yaacovi et al., 2013; see also, Jung & Vartanian, 2018). However, experimental designs still vary considerably across studies even within specific domains, which may reflect a lack of clarity of how to assess creativity most effectively in neuroscience research.

The valid assessment of creativity is already a big challenge in behavioral creativity research. All the more, this applies to cognitive neuroscience, because neurophysiological assessments impose additional methodological constraints (Abraham, 2013; Sawyer, 2011). Brain research requires the measurement of brain activation during well-defined cognitive activities that need to be repeatable over dozens of times in order to obtain reliable assessments. Moreover, neurophysiological assessments are sensitive to motor artifacts (e.g., caused by speaking, moving, or even blinking), which seriously limits the feasible ways of creative expression. Finally, neuroscientific investigations often involve test settings that appear not very conducive to creative thought (e.g. lying supine in a rattling MRI scanner or being wired with EEG electrodes). This general situation is quite at odds with the assessment of creative cognition, which typically represents complex cognitive activities that can take minutes and often require the generation of elaborated ideas and products. Therefore, neuroscience research on creativity is challenged to adapt creativity assessments to meet the constraints imposed by neurophysiological measurements. This article provides a brief review of how methodological issues in this field have been addressed so far. It focuses on functional imaging studies that investigate brain activation during actual creative task performance. The article hence does not cover other important lines of neuroscience research such as structural imaging, resting-state analyses, or lesion studies that measure creativity independent from brain assessments. These latter approaches face the same psychometric challenges as behavioral research on creativity, which are addressed by other articles in this Special Issue. We discuss the pros and cons associated with different methodological approaches, and derive some recommendations on how creativity can be studied in a reliable and valid way in cognitive neuroscience research.

Reviewing common practices in neuroscience research on creativity

General overview

To obtain a comprehensive picture of how neuroscience studies have approached creativity so far, we performed a quantitative review of available research in this field. We were interested in studies that investigated brain activity during creative cognition using common neuroscience methods including functional magnetic resonance imaging (fMRI), electroencephalography (EEG), near-infrared spectroscopy (NIRS), positron emission tomography (PET), and magnetoencephalography (MEG). Specifically, we searched the literature using *Web of Science*™ for English empirical articles employing a topic search considering title, abstract, and keywords with the search string: *TS = ((creativity OR "creative cognition" OR "creative thinking") AND (fMRI OR EEG OR MEG OR NIRS OR PET))*. Note that we used method-specific search cues instead of more general cues such as “brain” or “neuroscience” because it resulted in a more focused search of empirical work. This approach should provide a broad, representative overview of relevant work, but is not a fully exhaustive search, as it may miss studies that used relevant tasks, but did not mention "creativity/creative cognition/creative thinking" in the title, abstract or keywords. The search yielded 305 papers on October, 2nd, 2018. We excluded papers that included no original data (e.g., reviews), did not involve creative thinking tasks (e.g., metaphor comprehension), or did not study brain activity during actual creative task performance (e.g., structural MRI studies). This procedure left 115 articles describing a total of 131 different studies or tasks (some articles had multiple studies or tasks), which represented the final data considered in our analyses (data and scripts are provided at https://osf.io/zfr7v/).

An analysis of publications per year shows that, besides the pioneering work by Colin Martindale carried out in the 1970's and 1980's, neuroscience research on creativity didn't gain momentum until the end of 2000, with publication numbers increasing steadily ever since, and about 70% of articles having been published since 2010 (see Figure 1). From the studies included in this review, 48.9% used EEG, 48.1% used fMRI, and two studies each used PET and NIRS, respectively. While fMRI has only been used since 2005 in this field, it has outnumbered EEG in recent research and strongly contributes to the overall trend of increasing neuroscience research on creativity. Only one article combined two neuroimaging methods (i.e., EEG and fMRI; see Fink et al., 2009), suggesting a need for more multimodal research. The sample sizes in these studies range from 7 to 250, with an average of 35.81 (*SD* = 36.41; *median* = 28.0).

In the next step, we examined the prevalence of different kinds of creativity assessments in neuroscience research. We found that the majority of studies (51.1%) used divergent thinking (DT) tasks, which require to generate creative ideas to open-ended problems (e.g., alternate uses task); 19.1% of studies employed creative problem solving (CPS) tasks, which have correct solutions and often require a restructuring of the problem representation (e.g., Remote Associates Test, insight tasks); 29.8% of studies used a variety of mostly product-based tasks, which assess creative performances resulting in a creative product such as drawing, writing, or musical improvisation. These rates are very similar to those obtained in a recent review of behavioral creativity research based on a random stratified sample of 200 articles from 2009-2012 (Forgeard & Kaufman, 2016):

from 81 studies that directly assessed creative performance, 59.3% employed DT tasks (traditional or complex), 9.9% employed CPS tasks, and 30.9% used other tasks rated with the consensual assessment technique. This congruence (including a similar focus on DT tasks) suggests that the popularity of these task types in the behavioral research tradition is mirrored in neuroscience research.

We further analyzed the relative frequency of task domains (e.g., verbal, visual, music) defined by the type of response in these tasks. This analysis revealed that most of the employed creativity tasks collected verbal responses (73.3%), 17.6% were visual tasks (e.g., drawing), 7.6% were musical tasks, and 1.5% used other or mixed modalities (e.g., freestyle rap). It needs to be noted that the large set of verbal tasks was not homogenous. It included tasks that required verbal creativity in a narrower sense, for example, metaphor generation or story writing, but also many other tasks that just required a verbal response but without involving verbal creativity to the same degree. The alternate uses task, for example, has been commonly labeled as a verbal task (Torrance, 1974), although it may not require much verbal creativity to find creative object uses (Benedek, Fink & Neubauer, 2006). Only after an idea is generated, some verbal creativity may be involved in how well an idea is sold (Forthmann et al., 2017), suggesting that even different modalities can be prevalent at different stages of the task. In general, verbal responses are arguably the most convenient way to communicate abstract ideas. We hence believe that the prevalence of verbal creativity is overestimated in the literature, and this variability in task classification may contribute to inconsistencies in research findings. Therefore, traditional approaches to classify task domains by the modality of responses need to be reconsidered and shifted towards a focus on the modality of cognitive representations during the actual task.

How to adapt creativity tasks in neuroscience research?

Number of tasks and time on task. Neuroscience scholars commonly seek to employ proved and tested tasks similar to how they are used in behavioral research, but several adaptions are needed to meet the requirements of neurophysiological measurements. First of all, tasks are commonly embedded in more extended trials that can include task cues (i.e., reminding participants of the current task and condition, since thorough task instructions are often administered before entering the scanner or mounting EEG electrodes), reference periods (i.e., measurements of baseline brain activity, which often ask participants to relax while looking at a fixation cross), the actual task, and a response period. Moreover, neuroscience assessments typically involve many task repetitions (i.e., trials) and shorter task durations than behavioral assessments. The main reason for this is that neuroscience assessments do not target the measurement of a latent ability but rather the assessment of brain processes related to creative task performance. Establishing the relationship between creative cognition and brain activity involves several assumptions that are sometimes not well specified. First of all, we need to define what cognitive processes are involved when in the creative task. Yet, creative thinking tasks are often quite complex and involve many different processes and strategies that may vary over time (e.g., Gilhooly, Fioratou, Anthony & Wynn, 2007). Second, we need to define how cognitive processes manifest in a neurophysiological response. Different neurophysiological responses are commonly considered, including event-related potentials (ERP), and shifts in oscillatory activity in the EEG, or blood-oxygen-level dependent (BOLD) responses in fMRI assessments, and different responses follow different temporal dynamics. Crucially, neural responses of cognitive processes cannot be observed in isolation, but other ongoing processes as well as measurement artifacts (e.g. motion-related artifacts) add noise to the signal we are interested in. Further noise is contributed by inter- and intraindividual differences in cognitive processes and brain responses, which are typically hard to account for (Grady & Garrett, 2014). Given these potential sources of measurement error, many task repetitions (i.e., trials) are needed to ensure sufficient reliability of neurophysiological assessments. A large number of trials can only be realized with short tasks, which implies considerable shortening for extended creativity tasks (e.g. divergent thinking tasks). Shorter tasks may additionally reduce the variability of cognitive processes within tasks and better conform to neurophysiological models of brain responses.

Table 1 presents a descriptive analysis of the number of tasks and task duration (defined as the time needed to generate a solution or product, which excludes any time periods designated for instruction or response elaboration) in relevant research. Across all studies, the median number of tasks (i.e., trials) was 15 and the median task duration was 30 s. Notably, the correlation between number and duration of tasks showed a strong negative correlation (r_s = -.75), indicating that studies using shorter tasks afforded substantially more tasks. Task number and duration vary substantially across studies and differ between imaging methods and task type. For example, fMRI studies generally used more tasks (*median* = 20) than EEG studies (*median* = 3), but task duration was shorter in fMRI studies (*median* = 15 s) compared to EEG studies (*median* = 165 s). In part, this difference may be due to the fact that EEG assessments involve a less obtrusive assessment setting, so that tasks can be implemented more akin to standard cognitive testing. Moreover, EEG studies often target at the analysis of changes in brain activation over time within tasks (e.g., Jung-Beeman et al., 2004; Schwab, Benedek, Papousek, Weiss, & Fink, 2012). Studies on creative problem solving have used the highest number of tasks, followed by divergent thinking and other tasks, whereas time on task was typically higher for divergent thinking tasks compared to creative problem solving. These findings reflect differences in task duration that are also observed in behavioral testing (e.g., creative problem solving tasks are often realized with shorter time on task than divergent thinking tasks). Some studies even used as few as one task, which may have been done to establish basic brain activity or connectivity patterns for a complex task (e.g., De Pisapia, Bacci, Parrott, & Melcher, 2016). Some of the task durations are also very long (from 30 s to several minutes), which typically reflects self-paced tasks that require ongoing production such as drawing, musical improvisation, or tasks where multiple responses can be given within the task, which were later split into separate trials.

Our analysis shows that neuroscience studies on creativity typically employ around 15 tasks (i.e., trials) per condition or task type. This is much more than the number of tasks commonly employed in psychometric testing of the same constructs, but probably still less than the average number of tasks used in other fields of cognitive neuroscience (Abraham, 2013). The adequate number of tasks likely depends on the signal-to-noise ratio of the task and further design characteristics such as task duration and event-related vs. blocked task presentation (Friston, 1999). Moreover, for fMRI research, power calculation tools are available to determine the required sample size for a given design (Mumford & Nichols, 2008). High task numbers are only possible with short tasks, which comes with certain limitations. For example, divergent thinking tasks probably cannot be much shorter than 10-15 s to enable at least one valid response. Task periods of 15 s easily add up to a trial duration of 45 s, including time for cues, response periods, and fixation periods. If we have just two conditions with 20 trials each, we already end up with a test session of 30 minutes. In fMRI studies, we often want to add short structural scans and include time for initial calibration, which easily reaches the maximum time that participants can stay engaged in a task and be exposed to MRI measurements (i.e., usually about 45 minutes). Tasks focusing on more elementary cognitive aspects underlying creative thought, like passive conceptual expansion during evaluation (Rutter et al., 2012) or association processes (Green et al., 2015), allow shorter task times and thus higher amounts of trials. Hence, increasing the number of trials is limited by the minimum amount of time needed to perform tasks and the maximum total session time.

Besides these technical restrictions to tweaking tasks for neuroscience research, there are also important conceptual issues to consider. How does the shortening of tasks affect task validity? In DT tasks, for example, creativity of responses tend to increase with time as early responses are often retrieved from memory and thus represent more common, less creative ideas (Beaty & Silvia, 2012; Gilhooly et al., 2007). Yet, there is evidence suggesting that, while reliability increases with time on task and number of responses, this may not equally apply to validity; in fact, task validity may even decrease when DT tasks get too long (Benedek, Mühlmann, Jauk, & Neubauer, 2013). Moreover, we still know little about how cognitive task demands change when tasks become more speeded (i.e., shorter task durations). One recent study showed that more speeded tasks (e.g., 2 minutes of divergent thinking) did not rely more on mental speed (Gs) than “unspeeded” tasks (e.g., 8 minutes of divergent thinking; Forthmann, Lips, Szardenings, Scharfen, & Holling, in press), but task durations can be considerably shorter than that in neuroscience research. A few studies have reported relationships between task performance during short tasks used in neuroscience assessments and external criteria. For example, Perchtold and colleagues (2018) found a high correlation between rated divergent thinking performance in 15 s tasks administered inside the scanner and 3 min tasks administered outside the scanner (*r* = .64). Similarly, unpublished latent variable analyses from Beaty and colleagues (2018) found a strong latent correlation of *r* = .61 between divergent thinking ability assessed with 12 s scanner tasks (people generated 23 single uses for 23 different objects) and 3 min lab tasks (people generated many creative uses for two objects), which again strongly correlated with self-reported creative behavior and accomplishments in the arts and sciences (scanner *r* = .50 and lab *r* = .32), providing validity evidence consistent with, and potentially even higher than, lab-based measures (e.g., Jauk, Benedek & Neubauer 2014).

Collecting responses. Another aspect that requires careful consideration when adapting creativity tasks to neuroscientific assessments is how and when responses are collected. Since neurophysiological measurements are sensitive to movement artifacts, measurements of brain activity should not get confounded with response-related motor activity. While subtle motor activity related to button presses is typically not considered a problem, creative products often require more complex response forms such as speaking, writing, drawing, or even playing an instrument. Artifact detection and correction tools do an increasingly good job to remove motor-related artifacts from the data, yet there are technical limits to restoring original data. Therefore, researchers typically ensure to separate creative thought from response production in time. This approach is also consistent with the aim to distinguish different phases in the creative process (e.g., generation, elaboration, evaluation; Barbot, 2018; Ellamil et al., 2012; Fink et al., 2018; Jankowska et al., 2018; Loesche et al., 2018; Rominger et al., 2018).

There are different ways to deal with this issue. First, we can have participants perform the task self-paced, thus allowing them to give their responses whenever they arise, and only afterwards individually classify times prior to the response as creative thinking time versus times after response onset as response production time. Adapted to neuroscientific assessments this could mean that participants press a button when they have an idea, then vocalize their idea, and press another button when they are done with the response, thereby providing time stamps for thinking and response periods (e.g., Boot et al., 2017; Fink, Benedek, Grabner, Staudt, & Neubauer, 2007). After a response, they may move on to the next task, or, if multiple responses are required, continue with the task until timeout. Another option is to use voice key analyses applied to audio recordings during task performance to obtain relevant timings (e.g., Benedek et al., 2014b). Yet another approach is to define fixed time intervals a priori for idea generation and response generation, respectively (e.g., Fink et al., 2009). Some researchers additionally ask participants to indicate the occurrence of solutions by button-presses during task performance with fixed timing (e.g., Heinonen et al., 2018). Finally, for some complex creative activities, response production cannot be easily delayed and brain activation is measured concurrent to the actual production of responses (e.g. musical improvisation or drawing; e.g., Pinho et al., 2016; Saggar et al., 2018). Such studies typically employ control tasks that involve highly similar response productions in order to limit the risk of systematic confounds by motor activity. Another question relates to when and how responses are collected. Some studies collected verbal (Camarda et al., 2018) or written responses (e.g., Erhard, Kessler, Neumann, Ortheil & Lotze, 2014), while others ask participants to draw (e.g., Saggar et al., 2018), or play piano using MR-compatible instruments (Pinho et al., 2016). Yet other studies ask for button presses to indicate that a solution was found, which is often followed up by asking participants to recall their responses after the session or by assessing independent performance measures (e.g., Abraham et al., 2012; Vartanian et al., 2018).

Table 2 presents an analysis of the typical response timing (i.e., when are responses collected) and response type (i.e., what type of responses are collected) across neuroscience studies on creativity. Most commonly, studies have defined a fixed timing for thinking and response periods, either without button presses (40.1%) or with additional button presses during task performance to measure response times (16.3%). This method is especially popular in fMRI studies (73.0% in total). It is common across all types of tasks, but creative problem solving studies have used fixed timing more often together with additional button press (52.5%) and other tasks more often without. 29.3% of studies employed a self-paced approach, where the separation between thinking and response periods is realized post hoc relative to individual response onsets. This approach is especially common in EEG research (47.4%). Only 12.2% of studies have collected responses during the task, thus not separating thinking periods from response periods.

The self-paced approach aims to implement tasks as similar as possible to how they are typically administered in psychometric tests, allowing participants to respond whenever they come to a solution (Fink et al., 2007). We can assume that this approach achieves a validity similar to performance in standard cognitive testing. As a potential downside, this approach results in a variable amount of time spent on the task across participants (if the task ends as soon as a response is given; e.g., Tik et al., in press) or in a variable number of trials (if participants can give several responses within each task; e.g., Boot et al., 2017). When a variable amount of data is collected per participant this entails differences in the reliability of assessments, and it may even imply a direct confound, as more creative people typically respond more fluently. Following this approach, experimenters should make sure to obtain a certain minimum amount of trials/time for each participant.

The fixed-timing approach ensures that the number of task repetitions (i.e., trials) and time spent per task is equal across participants, which implies higher experimental control and also enables a more straightforward analysis of brain data. As a potential downside of this approach, however, it is not easy to specify a timing that works well for all task conditions and participants – it should be long enough to capture the essential process and provide all participants enough time to come up with a response, yet it should not be too long, as this can result in idling when participants find a response early in the task. The latter issue is commonly dealt with by asking participants to continue searching for even more creative ideas or to continue adding details to their response until the predefined time is over (e.g., Beaty, Silvia, Benedek, 2017; Benedek et al., 2014a).

Neuroscience studies also differ considerably in what kind of responses are collected. Most commonly they relied on oral responses (46.4%), but sometimes they also collected written responses or drawings, and about 25% of studies acquired button presses. Oral responses are most frequently collected in EEG research (56.7%) but also common in fMRI studies (33.9%). The review further shows that oral responses are most common in divergent thinking research (66.2%), whereas creative problem solving research more often use button presses (58.3%). These rates are certainly different from other fields of cognitive neuroscience, which mostly rely on button press responses. These findings highlight that most studies make an effort to collect responses in a way that pays tribute to the complex nature of creative productions—that is, creative performance cannot easily be reduced to button presses and typically involves speaking, writing, drawing or other formats to adequately communicate creative ideas and products.

Collecting responses in neuroscience of creativity often requires some extra effort (e.g., employing an MRI-compatible microphone), but is important for several reasons. First of all, it allows the researcher to monitor task performance and thus, substantiate that participants have been properly engaged in the task. Related to this, it facilitates the researcher's ability to discard trials from analysis where participants failed to follow instructions or to come up with a response. This is common practice in cognitive neuroscience research, but only possible when responses were collected in the first place. From the perspective of the participant, tasks may be more engaging when they require a response, whereas doing tasks just mentally can be tedious and increase the risk for mind wandering. As another important benefit, responses can be scored to obtain measures of individual performance. Scoring creative performance typically involves evaluations by several raters that show reasonable inter-rater-reliability. Performance data can serve to run manipulation checks (e.g., how do experimental conditions affect creative performance?; Benedek et al., 2018; Fink et al., 2012), and they enable additional lines of analyses relating task performance to brain activation. This can be done at the within-subjects level (i.e., how does brain activity differ between more vs. less original ideas?; e.g., Green et al., 2015) as well as at the between-subjects level (i.e., how does brain activity differ between more creative people vs. less creative people, as defined by task performance; e.g., Fink & Neubauer, 2008). Finally, scored task performance can be related to relevant external criteria in order to demonstrate the validity of adapted creativity tasks (Beaty et al., 2018). In sum, assessing creative performance during neurophysiological assessments is challenging but crucial to achieve a stronger inference on the actual relationship between creativity and brain activation.

Summary and conclusions

The investigation of the neural correlates of creative cognition requires researchers to adapt creativity tasks to meet the constraints imposed by cognitive neuroscience research – assessing well-defined cognitive processes, repeated over many tasks. A review of available neuroscience research of creativity revealed a large variability in essential study design parameters such as number of tasks, time on task, and when and what kind of responses are collected. In general, the employed creativity tasks are strongly inspired by available psychometric tests of creative potential that were adapted towards lower time on task to allow running more tasks per study. Most studies have collected qualitative responses such as oral, written, drawn or musical productions, which enable to analyze brain activation related to creative performance. Moreover, studies usually take care to separate times of creative thought from response production in order to limit potential confounds with response-related motor activity. Taken together these adaptations are useful to increase reliability of neurophysiological measurements while maintaining valid assessments. Indeed, initial evidence suggests that divergent thinking tasks that are adapted for neuroscience, for example, perform well in terms of concurrent validity as well as criterion validity (Beaty et al., 2018; Perchtold et al., 2018).

While the development of neuroscience paradigms has strongly built on paradigms from behavioral creativity research, we believe that neuroscience research also bears large potential to inform behavioral research. First, neuroscience research requires great rigor regarding assumptions on the involvement and timing of relevant cognitive processes. This has fueled new interest in the examination of specific attention and memory processes in creative thought. For example, the robust association between creativity and EEG alpha activity (Fink & Benedek, 2014; Kounios & Beeman, 2014) inspired cognitive research on internally directed attention in creative cognition (e.g., Ritter et al., 2018; Salvi & Bowden, 2016; Walcher et al., 2017; for a review, see Benedek, 2018). Or, the persistent relevance of the default network in creative cognition (Beaty et al., 2016; Zabelina & Andrews-Hanna, 2016) attracted much interest in the role of episodic memory for creative thought (e.g., Madore et al., 2016). Second, the constraints imposed by neuroscience research have stimulated novel types and variants of creativity tasks. These tasks are often very well defined in terms of their cognitive demands (Barbot, 2018; Prabhakaran, Green & Gray, 2014) and show promising psychometric quality (e.g., Beaty et al., 2018). Hence, addressing the challenges for assessment of creativity in neuroscience research has inspired creative solutions to creativity assessment that may turn out to be more than just purposeful adaptations.

We conclude with some recommendations for future research. First, of course, it is crucial to employ tasks that capture relevant aspects of creative cognition, thereby assuring optimal validity. The demonstration of validity evidence (e.g., correlations with performance in original tasks or external criteria of creativity) is essential for novel creativity tasks but is also recommended for established tasks that have undergone substantial adaptations for neuroscience assessments. Second, cognitive neuroscience research requires clear a priori assumptions on what cognitive processes interact in creative tasks. Neuroscientific assessments of highly complex artistic performances are intriguing, but they typically do not allow to reliably relate brain activation to specific psychological processes and thus heavily rely on reverse inference (Poldrack, 2006). Powerful tests of brain-cognition associations need to specify what cognitive processes (e.g., memory, attention, cognitive control processes) are central to the main task, how they differ in control tasks (but see Logothetis, 2008, for a discussion of pure insertion issues), and hypothesize on causal relationships between neurocognitive processes (e.g., Vartanian et al., 2018). To this end, the design of neurophysiological assessments needs to be guided by and based on the rich evidence of cognitive science. Third, disentangling cognitive processes may further imply to distinguish specific phases or stages in the creative process as presumed by available theoretical models (e.g., generation, evaluation, elaboration). Fourth, creative cognition stands out in that it commonly involves the generation of ideas or products that differ in quality. Assessing these productions and their creative quality rather than just the time of their occurrence enables powerful analyses of creativity-related brain functions. Fifth, co-registration studies assessing different neurophysiological parameters concurrently are needed to relate and consolidate evidence across neuroscience methods. Finally, as all fields of cognitive neurosciences, the neuroscience of creativity needs to ensure that studies are well-powered in terms of sufficient trials and sample size (Yarkoni, 2009). In this manner, task-based creativity neuroscience, together with other techniques including structural imaging (Jung et al., 2013), brain stimulation (Weinberger et al., 2017), and neuropsychological approaches (Abraham, 2018) will help us to advance our understanding of how creativity emerges in the brain.

References

Abraham, A. (2018). The neuropsychology of creativity. *Current Opinions in Behavioral Sciences, 27*, 71-76.

Abraham, A. (2018). *The neuroscience of creativity*. Cambridge: Cambridge University Press.

Abraham, A. (2013). The promises and perils of the neuroscience of creativity. *Frontiers in Human Neuroscience, 7*, 246.

Abraham, A., Pieritz, K., Thybusch, K., Rutter, B., Kröger, S., Schweckendiek, J., ... & Hermann, C. (2012). Creativity and the brain: uncovering the neural signature of conceptual expansion. *Neuropsychologia, 50*, 1906-1917.

Arden, R., Chavez, R. S., Grazioplene, R., & Jung, R. E. (2010). Neuroimaging creativity: a psychometric view. *Behavioural Brain Research, 214*, 143-156.

Barbot, B. (2018). Measuring the dynamics of creative ideation: A new assessment paradigm. Presentation at the Meeting of the Society for the Neuroscience of Creativity. Cambridge, MA.

Beaty, R.E., Benedek, M., Silvia, P.J., & Schacter, D.L. (2016). Creative cognition and brain network dynamics. *Trends in Cognitive Sciences, 20*, 87-95.

Beaty, R.E., Kenett, Y.N., Christensen, A.P., Rosenberg, M.D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T., Kane, M.J., & Silvia, P.J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences (PNAS), 115*, 1087-1092.

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 309-319.

Beaty, R.E., Silvia, P.J., & Benedek, M. (2017). Brain networks underlying novel metaphor production. *Brain and Cognition, 111*, 163-170.

Benedek, M. (2018). Internally directed attention in creative cognition. In R. E. Jung & O. Vartanian (Eds.), *The Cambridge handbook of the neuroscience of creativity* (pp. 180-194). New York: Cambridge University Press.

Benedek, M., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., Silvia, P. J., Dunst, B., & Neubauer, A.C. (2014a). Creating metaphors: The neural basis of figurative language production. *NeuroImage, 90*, 99-106.

Benedek, M., Fink, A. & Neubauer, A.C. (2006). Enhancement of ideational fluency by means of computer-based training. *Creativity Research Journal, 18*, 317-328.

Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014b). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *NeuroImage, 88*, 125-133.

Benedek, M., Mühlmann, C., Jauk, E., & Neubauer, A. C. (2013). Assessment of divergent thinking by means of the subjective top-scoring method: Effects of the number of top-ideas and time-on-task on reliability and validity. *Psychology of Aesthetics, Creativity, and the Arts, 7*, 341-349.

Benedek, M., Schües, T., Beaty, R., Jauk, E., Koschutnig, K., Fink, A., & Neubauer, A. C. (2018) To create or to recall original ideas: Brain processes associated with the imagination of novel object uses. *Cortex, 99*, 93-102.

Boot, N., Baas, M., Mühlfeld, E., de Dre, C.W.K., & van Gaal, S. (2017). Widespread neural oscillations in the delta band dissociate rule convergence from rule divergence during creative idea generation. *Neuropsychologia, 104*, 8-17.

Camarda, A., Salvia, E., Vidal, J., Weil, B., Poirel, N., Houde, O., ... & Cassotti, M. (2018). Neural basis of functional fixedness during creative idea generation: an EEG study. *Neuropsychologia, 188*, 4-12.

De Pisapia, N., Bacci, F., Parrott, D., & Melcher, D. (2016). Brain networks for visual creativity: a functional connectivity study of planning a visual artwork. *Scientific reports, 6*, 39185.

Dietrich, A. (Eds.) (2007). Neurocognitive Mechanisms of Creativity: A Toolkit [Special Issue]. *Methods, 42*(1).

Ellamil, M., Dobson, C., Beemna, M., & Christoff, K. (2012). Evaluative and generative modes of thought during the creative process. *NeuroImage, 59*, 1783-1794.

Erhard, K., Kessler, F., Neumann, N., Ortheil, H. J., & Lotze, M. (2014). Professional training in creative writing is associated with enhanced fronto-striatal activity in a literary text continuation task. *NeuroImage, 100*, 15-23.

Fink, A., & Benedek, M. (2014). EEG alpha power and creative ideation. *Neuroscience and Biobehavioral Reviews, 44*, 111-123.

Fink, A., Benedek, M., Grabner, R. H., Staudt, B. & Neubauer, A.C. (2007). Creativity meets neuroscience: Experimental tasks for the neuroscientific study of creative thinking. *Methods, 42*, 67-75.

Fink, A., Grabner, R. H., Benedek, M., Reishofer, G., Hauswirth, V., Fally, M., Neuper, C., Ebner, F. & Neubauer, A.C. (2009). The creative brain: Investigation of brain activity during creative problem solving by means of EEG and fMRI. *Human Brain Mapping, 30*, 734-748.

Fink, A., Koschutnig, K., Benedek, M., Reishofer, G., Ischebeck, A., Weiss, E. M., & Ebner, F. (2012). Stimulating creativity via exposure to other people's ideas. *Human Brain Mapping, 33*, 2603-2610.

Fink, A., & Neubauer, A. C. (2008). Eysenck meets Martindale: The relationship between extraversion and originality from the neuroscientific perspective. *Personality and Individual Differences, 44*, 299-310.

Fink, A., Rominger, C., Benedek, M., Perchtold, C., Papousek, I., Weiss, E.M., Seidel, A., & Memmert, D. (2018). EEG alpha activity during imagining creative moves in soccer decision-making situations. *Neuropsychologia, 114*, 118-124.

Foregeard, M. J. C., & Kaufman, J. C. (2016). Who cares about imagination, creativity, and innovation and why? A Review. *Psychology of Aesthetics, Creativity, and the Arts, 10*, 250-269.

Forthmann, B., Holling, H., Çelik, P., Storme, M., & Lubart, T. (2017). Typing speed as a confounding variable and the measurement of quality in divergent thinking. *Creativity Research Journal, 29*, 257-269.

Forthmann, B., Lips, C., Szardenings, C., Scharfen, J., & Holling, H. (in press.) Are speedy brains needed when divergent thinking is speeded - or unspeeded? *The Journal of Creative Behavior*. https://doi.org/10.1002/jocb.350

Gilhooly, K.J., Fioratou, E., Anthony, S.H., Wynn, V. (2007). Divergent thinking: strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98*, 611–625.

Grady, C. L., & Garrett, D. D. (2014). Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging and Behavior, 8*, 274-283.

Green, A. E., Cohen, M. S., Raab, H. A., Yedibalian, C. G., & Gray, J. R. (2015). Frontopolar activity and connectivity support dynamic conscious augmentation of creative state. *Human brain mapping, 36*, 923-934.

Gonen-Yaacovi, G., De Souza, L. C., Levy, R., Urbanski, M., Josse, G., & Volle, E. (2013). Rostral and caudal prefrontal contribution to creativity: a meta-analysis of functional imaging data. *Frontiers in Human Neuroscience, 7*, 465.

Heinonen, J., Numminen, J., Hlushchuk, Y., Antell, H., Taatila, V., & Suomala, J. (2016) Default mode and executive networks areas: Association with the serial order in divergent thinking. *PLoS ONE, 11*: e0162234.

Jankowska, D. M., Czerwonka, M., Lebuda, I., & Karwowski, M. (2018). Exploring the creative process:

Integrating psychometric and eye-tracking approaches. *Frontiers in Psychology, 9*, 1931.

Jauk, E., Benedek, M., & Neubauer, A. C. (2014). The road to creative achievement: A latent variable model of ability and personality predictors. *European Journal of Personality, 28*, 95-105.

Jung, R. E., Mead, B. S., Carrasco, J., & Flores, R. A. (2013). The structure of creative cognition in the human brain. *Frontiers in Human Neuroscience, 7*, 330.

Jung, R. E. & Vartanian O. (Eds.). (2018). The Cambridge handbook of the neuroscience of creativity. New York: Cambridge University Press.

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., ... & Kounios, J. (2004). Neural activity when people solve verbal problems with insight. *PLoS Biology, 2*, e97.

Kounios, J., & Beeman, M. (2014). The cognitive neuroscience of insight. *Annual Review of Psychology, 65*, 71-93.

Loesche, F., Goslin, J., & Bugmann, G. (2018). Paving the Way to Eureka—Introducing “Dira” as an Experimental Paradigm to Observe the Process of Creative Problem Solving. *Frontiers in Psychology, 9*, 1773.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*, 869-878.

Madore, K. P., Jin, H., G., & Schacter, D. L., (2016). Divergent creative thinking in young and older adults: Extending the effects of an episodic specificity induction. *Memory & Cognition, 44*, 974-988.

Perchtold, C. M., Papousek, I., Koschutnig, K., Rominger, C., Weber, H., Weiss, E. M., & Fink, A. (2018). Affective creativity meets classic creativity in the scanner. *Human Brain Mapping, 39*, 393-406.

Poldrack, R. A. (2006) Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*, 59-63.

Prabhakaran, R., Green, A. E., & Gray, J. R. (2014). Thin slices of creativity: Using single-word utterances to assess creative cognition. *Behavior Research Methods, 46*, 641-659.

Rominger, C., Papousek, I., Perchtold, C.M., Weber, B., Weiss, E.M., & Fink, A. (2018). The creative brain in the figural domain: Distinct patterns of EEG alpha power during idea generation and idea elaboration. *Neuropsychologia, 118*, 13-19.

Ritter, S.M., Abbing, J. & van Schie, H.T. (2018) Eye-closure enhances creative performance on divergent and convergent creativity tasks. *Frontiers in Psychology, 9*, 1315.

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (2012). Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion. *Brain and Cognition, 80*, 301-210.

Saggar, M., Quintin, E. M., Kienitz, E., Bott, N. T., Sun, Z., Hong, W. C., ... & Hawthorne, G. (2015). Pictionary-based fMRI paradigm to study the neural correlates of spontaneous improvisation and figural creativity. *Scientific Reports, 5*, 10894.

Salvi, C., & Bowden, E. M. (2016). Looking for creativity: Where do we look when we look for new ideas? *Frontiers in Psychology, 7*, 161.

Sawyer, K. (2011). The cognitive neuroscience of creativity: a critical review. *Creativity Research Journal, 23*, 137-154.

Schwab, D., Benedek, M., Papousek, I., Weiss, E.M., & Fink, A. (2014). The time-course of EEG alpha power changes in creative ideation. *Frontiers in Human Neuroscience, 8*, 310.

Tik, M., Sladky, R., Luft, C. D. B., Willinger, D., Hoffmann, A., Banissy, M. J., ... & Windischberger, C. (in press). Ultra-high-field fMRI insights on insight: Neural correlates of the Aha!-moment. *Human Brain Mapping*.Torrance, E. P. (1974). *Torrance Tests of Creative Thinking: Norms, technical manual, verbal forms A and B*. Bensenville, IL: Scholastic Testing Service.

Vartanian, O., Beatty, E. L., Smith, I., Blackler, K., Lam, Q., & Forbes, S. (2018). One-way traffic: The inferior frontal gyrus controls brain activation in the middle temporal gyrus and inferior parietal lobule during divergent thinking. *Neuropsychologia, 118*, 68-78..Walcher, S., Körner, C., & Benedek, M. (2017). Looking for ideas: Eye behavior during goal-directed internally-focused cognition. *Consciousness and Cognition, 53*, 165-175.

Weinberger, A. B., Green, A. E., & Chrysikou, E. G. (2017). Using transcranial direct current stimulation to enhance creative cognition: interactions between task, polarity, and stimulation site. *Frontiers in Human Neuroscience, 11*, 246.

Yarkoni, T. (2009). Big correlations in little studies: Inflated fMRI correlations reflect low statistical power—Commentary on Vul et al. (2009). *Perspectives on Psychological Science, 4*, 294-298.

Zabelina, D. L., & Andrews-Hanna, J. R. (2016). Dynamic network interactions supporting internally-oriented cognition. *Current Opinion in Neurobiology, 40*, 86-93.

TABLES

Table 1

Analysis of the number of tasks (i.e., trials) and task duration [s] in neuroscience studies on creativity (1975-2018). Descriptive statistics for all studies, and separate for EEG and fMRI studies, and studies on divergent thinking (DT), creative problem solving (CPS), or other tasks (i.e., mostly product-based tasks such as creative writing, drawing or musical improvisation).

# Tasks	*n*	*Median*	*M*	*SD*	*Min*	*Max*
All studies	117	15	24.53	33.74	1	210
EEG	53	3	12.62	26.07	1	150
fMRI	61	20	35.87	36.54	1	210
DT	60	13.5	15.63	14.76	1	72
CPS	22	33	42.82	40.65	1	150
Other tasks	35	10	28.29	45.98	1	210

Time on task [s]	*n*	*Median*	*M*	*SD*	*Min*	*Max*
All studies	117	30	132.53	434.91	2	4500
EEG	54	165	242.91	616.95	4	4500
fMRI	60	15	36.81	95.89	2	600
DT	62	30	154.24	576.27	2	4500
CPS	20	19	140.18	231.00	4.5	900
Other tasks	35	30	89.69	130.02	2	480

Note. *n* represents the number of valid data in each analysis, after excluding studies that failed to report relevant information.

Table 2

Analysis of when and what kind of responses are collected in neuroscience studies on creativity (1975-2018). Descriptive statistics for all studies, and separate for EEG and fMRI studies, and studies on divergent thinking (DT), creative problem solving (CPS), or other tasks.

Response timing [%]	*n*	After task (Self-paced)	After task (Fixed)	After task (Fixed + RT)	During task	No response
All studies	123	29.3	40.1	16.3	12.2	1.6
EEG	57	47.4	33.3	5.3	12.3	1.8
fMRI	63	12.7	46.0	27.0	12.7	1.6
DT	64	35.9	48.4	6.3	7.8	1.6
CPS	23	34.8	13.0	52.2	0.0	0.0
Other tasks	36	13.9	44.4	11.1	27.8	2.8

Response type [%]		Speak	Write	Draw	Button	Other	None
All studies	125	46.4	12.0	12.8	23.2	4.8	0.8
EEG	60	56.7	18.3	16.7	5.0	1.7	1.7
fMRI	62	33.9	6.5	9.7	41.9	8.1	0.0
DT	65	66.2	13.8	7.7	12.3	0.0	0.0
CPS	24	33.3	8.3	0.0	58.3	0.0	0.0
Other tasks	36	19.4	11.1	30.6	19.4	16.7	2.8

Notes. *n* represents the number of valid data in each analysis, after excluding studies that failed to report relevant information. After task (self-paced) = participants can give their response at any time, with the response onset marking the end of generation period and the beginning of the response period; after task (fixed) = fixed thinking time and response time were specified by the experimenter; after task (fixed + RT) = fixed thinking time and response time were specified by the experimenter, but participant gives button press when a solution occurs; during task = responses were collected concurrently to task performance.

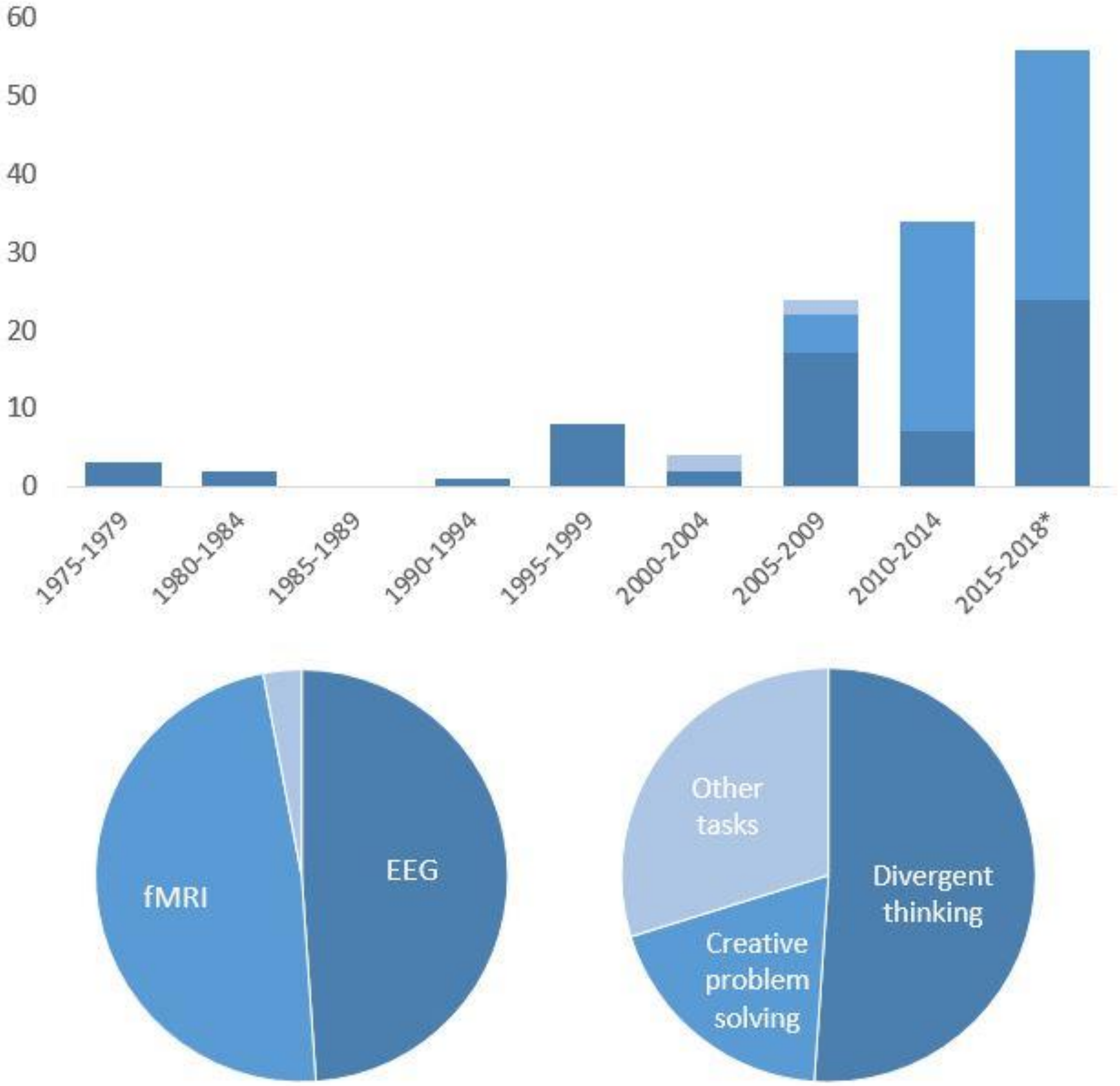

Figure 1. Top: Number of articles on the neuroscience of creativity over time per method (dark blue = EEG, blue = fMRI, light blue = other method). *Note that time bins represent 5 years, except for the last, most recent bin, which only covers a little more than 3 years, i.e., 01/2015-10/2018. Bottom left: Relative frequency of neuroscience studies employing fMRI, EEG, or other methods. Bottom right: Relative frequency of neuroscience studies investigating divergent thinking, creative problem solving, or other tasks.